Big Data Et Machine Learning Manuel Du Data Scientist Management Des Systegravemes Dinformation

This book is intended to present the state of the art in research on machine learning and big data analytics. The accepted chapters covered many themes including artificial intelligence and data mining applications, machine learning and applications, deep learning technology for big data analytics, and modeling, simulation, and security with big data. It is a valuable resource for researchers in the area of big data analytics and its applications.

This book constitutes the post-conference proceedings of the Third International Workshop on Machine Learning, Optimization, and Big Data, MOD 2017, held in Volterra, Italy, in September 2017. The 50 full papers presented were carefully reviewed and selected from 126 submissions. The papers cover topics in the field of machine learning, artificial intelligence, computational optimization and data science presenting a substantial array of ideas, technologies, algorithms, methods and applications.

This book provides a comprehensive picture of fog computing technology, including of fog architectures, latency aware application management issues with real time requirements, security and privacy issues and fog analytics, in wide ranging application scenarios such as M2M device communication, smart homes, smart vehicles, augmented reality and transportation management. This book explores the research issues involved in the application of traditional shallow machine learning and deep learning techniques to big data analytics. It surveys global research advances in extending the conventional unsupervised or clustering algorithms, extending supervised and semi-supervised algorithms and association rule mining algorithms to big data Scenarios. Further it discusses the deep learning applications of big data analytics to fields of computer vision and speech processing, and describes applications such as semantic indexing and data tagging. Lastly it identifies 25 unsolved research problems and research directions in fog computing, as well as in the context of applying deep learning techniques to big data analytics, such as dimensionality reduction in high-dimensional data and improved formulation of data abstractions along with possible directions for their solutions.

Many approaches have sprouted from artificial intelligence (AI) and produced major breakthroughs in the computer science and engineering industries. Deep learning is a method that is transforming the world of data and analytics. Optimization of this new approach is still unclear, however, and there's a need for research on the various applications and techniques of deep learning in the field of computing. Deep Learning Techniques and Optimization Strategies in Big Data Analytics is a collection of innovative research on the methods and applications of deep learning strategies in the fields of computer science and information systems. While highlighting topics including data integration, computational modeling, and scheduling systems, this book is ideally designed for engineers, IT specialists, data analysts, data scientists, engineers, researchers, academicians, and students seeking current research on deep learning methods and its application in the digital industry.

This compendium provides a comprehensive collection of the emergent applications of big data, machine learning, and artificial intelligence technologies to present day physical sciences ranging from materials theory and imaging to predictive synthesis and automated research. This area of research is among the most rapidly developing in the last several years in areas spanning materials science, chemistry, and condensed matter physics.Written by world renowned researchers, the compilation of two authoritative volumes provides a distinct summary of the modern advances in instrument — driven data generation and analytics, establishing the links between the big data and predictive theories, and outlining the emerging field of data and physics-driven predictive and autonomous systems.

Cet ouvrage s'adresse à tous ceux qui cherchent à tirer parti de l'énorme potentiel des technologies Big Data, qu'ils soient data scientists, DSI, chefs de projets ou spécialistes métier. Le Big Data s'est imposé comme une innovation majeure pour toutes les entreprises qui cherchent à construire un avantage concurrentiel de l'exploitation de leurs données clients, fournisseurs, produits, processus, etc. Il a en outre permis l'émergence des techniques d'apprentissage automatique (Machine Learning, Deep Learning...) qui ont relancé le domaine de l'intelligence artificielle. Mais quelle solution technique choisir ? Quelles compétences métier développer au sein de la DSI ? Ce livre est un guide pour comprendre les enjeux d'un projet Big Data, en appréhender les concepts sous-jacents et acquérir les compétences nécessaires à la mise en place d'une architecture d'entreprise adaptée. Il combine la présentation : de notions théoriques (traitement statistique des données, calcul distribué...) ; des outils les plus répandus ; d'exemples d'applications, notamment en NLP (Natural Language Processing) ; d'une organisation typique d'un projet de data science.

Provides an up-to-date analysis of big data and multi-agent systems The term Big Data refers to the cases, where data sets are too large or too complex for traditional data-processing software. With the spread of new concepts such as Edge Computing or the Internet of Things, production, processing and consumption of this data becomes more and more distributed. As a result, applications increasingly require multiple agents that can work together. A multi-agent system (MAS) is a self-organized computer system that comprises multiple intelligent agents interacting to solve problems that are beyond the capacities of individual agents. Modern Big Data Architectures examines modern concepts and architecture for Big Data processing and analytics. This unique, up-to-date volume provides joint analysis of big data and multi-agent systems, with emphasis on distributed, intelligent processing of very large data sets. Each chapter contains practical examples and detailed solutions suitable for a wide variety of applications. The author, an internationally-recognized expert in Big Data and distributed Artificial Intelligence, demonstrates how base concepts such as agent, actor, and micro-service have reached a point of convergence—enabling next generation systems to be built by incorporating the best aspects of the field. This book: Illustrates how data sets are produced and how they can be utilized in various areas of industry and science Explains how to apply common computational models and state-of-the-art architectures to process Big Data tasks Discusses current and emerging Big Data applications of Artificial Intelligence Modern Big Data Architectures: A Multi-Agent Systems Perspective is a timely and important resource for data science professionals and students involved in Big Data analytics, and machine and artificial learning.

Offers a clear view of the utility and place for survey data within the broader Big Data ecosystem This book presents a collection of snapshots from two sides of the Big Data perspective. It assembles an array of tangible tools, methods, and approaches that illustrate how Big Data sources and methods are being used in the survey and social sciences to improve official statistics and estimates for human populations. It also provides examples of how survey data are being used to evaluate and improve the quality of insights derived from Big Data. Big Data Meets

Survey Science: A Collection of Innovative Methods shows how survey data and Big Data are used together for the benefit of one or more sources of data, with numerous chapters providing consistent illustrations and examples of survey data enriching the evaluation of Big Data sources. Examples of how machine learning, data mining, and other data science techniques are inserted into virtually every stage of the survey lifecycle are presented. Topics covered include: Total Error Frameworks for Found Data; Performance and Sensitivities of Home Detection on Mobile Phone Data; Assessing Community Wellbeing Using Google Street View and Satellite Imagery; Using Surveys to Build and Assess RBS Religious Flag; and more. Presents groundbreaking survey methods being utilized today in the field of Big Data Explores how machine learning methods can be applied to the design, collection, and analysis of social science data Filled with examples and illustrations that show how survey data benefits Big Data evaluation Covers methods and applications used in combining Big Data with survey statistics Examines regulations as well as ethical and privacy issues Big Data Meets Survey Science: A Collection of Innovative Methods is an excellent book for both the survey and social science communities as they learn to capitalize on this new revolution. It will also appeal to the broader data and computer science communities looking for new areas of application for emerging methods and data sources.

Big Data et Machine LearningLes concepts et les outils de la data science

Big Data Analytics in Cyber-Physical Systems: Machine Learning for the Internet of Things examines sensor signal processing, IoT gateways, optimization and decision-making, intelligent mobility, and implementation of machine learning algorithms in embedded systems. This book focuses on the interaction between IoT technology and the mathematical tools used to evaluate the extracted data of those systems. Each chapter provides the reader with a broad list of data analytics and machine learning methods for multiple IoT applications. Additionally, this volume addresses the educational transfer needed to incorporate these technologies into our society by examining new platforms for IoT in schools, new courses and concepts for universities and adult education on IoT and data science. . Bridges the gap between IoT, CPS, and mathematical modelling. Features numerous use cases that discuss how concepts are applied in different domains and applications. Provides "best practices", "winning stories" and "real-world examples" to complement innovation. Includes highlights of mathematical foundations of signal processing and machine learning in CPS and IoT.

"This textbook is a well-rounded, rigorous, and informative work presenting the mathematics behind modern machine learning techniques. It hits all the right notes: the choice of topics is up-to-date and perfect for a course on data science for mathematics students at the advanced undergraduate or early graduate level. This book fills a sorely-needed gap in the existing literature by not sacrificing depth for breadth, presenting proofs of major theorems and subsequent derivations, as well as providing a copious amount of Python code. I only wish a book like this had been around when I first began my journey!" -Nicholas Hoell, University of Toronto "This is a well-written book that provides a deeper dive into data-scientific methods than many introductory texts. The writing is clear, and the text logically builds up regularization, classification, and decision trees. Compared to its probable competitors, it carves out a unique niche. -Adam Loy, Carleton College The purpose of Data Science and Machine Learning: Mathematical and Statistical Methods is to provide an accessible, yet comprehensive textbook intended for students interested in gaining a better understanding of the mathematics and statistics that underpin the rich variety of ideas and machine learning algorithms in data science. Key Features: Focuses on mathematical understanding. Presentation is self-contained, accessible, and comprehensive. Extensive list of exercises and worked-out examples. Many concrete algorithms with Python code. Full color throughout. The Authors: Dirk P. Kroese, PhD, is a Professor of Mathematics and Statistics at The University of Queensland. He has published over 120 articles and five books in a wide range of areas in mathematics, statistics, data science, machine learning, and Monte Carlo methods. He is a pioneer of the well-known Cross-Entropy method—an adaptive Monte Carlo technique, which is being used around the world to help solve difficult estimation and optimization problems in science, engineering, and finance. Zdravko Botev, PhD, is an Australian Mathematical Science Institute Lecturer in Data Science and Machine Learning with an appointment at the University of New South Wales in Sydney, Australia. He is the recipient of the 2018 Christopher Heyde Medal of the Australian Academy of Science for distinguished research in the Mathematical Sciences. Thomas Taimre, PhD, is a Senior Lecturer of Mathematics and Statistics at The University of Queensland. His research interests range from applied probability and Monte Carlo methods to applied physics and the remarkably universal self-mixing effect in lasers. He has published over 100 articles, holds a patent, and is the coauthor of Handbook of Monte Carlo Methods (Wiley). Radislav Vaisman, PhD, is a Lecturer of Mathematics and Statistics at The University of Queensland. His research interests lie at the intersection of applied probability, machine learning, and computer science. He has published over 20 articles and two books.

This book contributes to the progress towards intelligent transportation. It emphasizes new data management and machine learning approaches such as big data, deep learning and reinforcement learning. Deep learning and big data are very energetic and vital research topics of today's technology. Road sensors, UAVs, GPS, CCTV and incident reports are sources of massive amount of data which are crucial to make serious traffic decisions. Herewith this substantial volume and velocity of data, it is challenging to build reliable prediction models based on machine learning methods and traditional relational database. Therefore, this book includes recent research works on big data, deep convolution networks and IoT-based smart solutions to limit the vehicle's speed in a particular region, to support autonomous safe driving and to detect animals on roads for mitigating animal-vehicle accidents. This book serves broad readers including researchers, academicians, students and working professional in vehicles manufacturing, health and transportation departments and networking companies.

The idea behind this book is to simplify the journey of aspiring readers and researchers to understand Big Data, IoT and Machine Learning. It also includes various real-

time/offline applications and case studies in the fields of engineering, computer science, information security and cloud computing using modern tools. This book consists of two sections: Section I contains the topics related to Applications of Machine Learning, and Section II addresses issues about Big Data, the Cloud and the Internet of Things. This brings all the related technologies into a single source so that undergraduate and postgraduate students, researchers, academicians and people in industry can easily understand them. Features Addresses the complete data science technologies workflow Explores basic and high-level concepts and services as a manual for those in the industry and at the same time can help beginners to understand both basic and advanced aspects of machine learning Covers data processing and security solutions in IoT and Big Data applications Offers adaptive, robust, scalable and reliable applications to develop solutions for day-to-day problems Presents security issues and data migration techniques of NoSQL databases

This book constitutes revised selected papers from the Second International Workshop on Machine Learning, Optimization, and Big Data, MOD 2016, held in Volterra, Italy, in August 2016. The 40 papers presented in this volume were carefully reviewed and selected from 97 submissions. These proceedings contain papers in the fields of Machine Learning, Computational Optimization and DataScience presenting a substantial array of ideas, technologies, algorithms, methods and applications.

This book presents the latest advances in machine intelligence and big data analytics to improve early warning of cyber-attacks, for cybersecurity intrusion detection and monitoring, and malware analysis. Cyber-attacks have posed real and wide-ranging threats for the information society. Detecting cyber-attacks becomes a challenge, not only because of the sophistication of attacks but also because of the large scale and complex nature of today's IT infrastructures. It discusses novel trends and achievements in machine intelligence and their role in the development of secure systems and identifies open and future research issues related to the application of machine intelligence in the cybersecurity field. Bridging an important gap between machine intelligence, big data, and cybersecurity communities, it aspires to provide a relevant reference for students, researchers, engineers, and professionals working in this area or those interested in grasping its diverse facets and exploring the latest advances on machine intelligence and big data analytics for cybersecurity applications.

This two-volume set, LNCS 12565 and 12566, constitutes the refereed proceedings of the 6th International Conference on Machine Learning, Optimization, and Data Science, LOD 2020, held in Siena, Italy, in July 2020. The total of 116 full papers presented in this two-volume post-conference proceedings set was carefully reviewed and selected from 209 submissions. These research articles were written by leading scientists in the fields of machine learning, artificial intelligence, reinforcement learning, computational optimization, and data science presenting a substantial array of ideas, technologies, algorithms, methods, and applications.

Summary Introducing Data Science teaches you how to accomplish the fundamental tasks that occupy data scientists. Using the Python language and common Python libraries, you'll experience firsthand the challenges of dealing with data at scale and gain a solid foundation in data science. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Many companies need developers with data science skills to work on projects ranging from social media marketing to machine learning. Discovering what you need to learn to begin a career as a data scientist can seem bewildering. This book is designed to help you get started. About the Book Introducing Data ScienceIntroducing Data Science explains vital data science concepts and teaches you how to accomplish the fundamental tasks that occupy data scientists. You'll explore data visualization, graph databases, the use of NoSQL, and the data science process. You'll use the Python language and common Python libraries as you experience firsthand the challenges of dealing with data at scale. Discover how Python allows you to gain insights from data sets so big that they need to be stored on multiple machines, or from data moving so quickly that no single machine can handle it. This book gives you hands-on experience with the most popular Python data science libraries, Scikit-learn and StatsModels. After reading this book, you'll have the solid foundation you need to start a career in data science. What's Inside Handling large data Introduction to machine learning Using Python to work with data Writing data science algorithms About the Reader This book assumes you're comfortable reading code in Python or a similar language, such as C, Ruby, or JavaScript. No prior experience with data science is required. About the Authors Davy Cielen, Arno D. B. Meysman, and Mohamed Ali are the founders and managing partners of Optimately and Maiton, where they focus on developing data science projects and solutions in various sectors. Table of Contents Data science in a big data world The data science process Machine learning Handling large data on a single computer First steps in big data Join the NoSQL movement The rise of graph databases Text mining and text analytics Data visualization to the end user

The ideas introduced in this book explore the relationships among rule based systems, machine learning and big data. Rule based systems are seen as a special type of expert systems, which can be built by using expert knowledge or learning from real data. The book focuses on the development and evaluation of rule based systems in terms of accuracy, efficiency and interpretability. In particular, a unified framework for building rule based systems, which consists of the operations of rule generation, rule simplification and rule representation, is presented. Each of these operations is detailed using specific methods or techniques. In addition, this book also presents some ensemble learning frameworks for building ensemble rule based systems. Now in its second edition, this book focuses on practical algorithms for mining data from even the largest datasets.

This book presents the original articles that have been accepted in the 2019 INNS Big Data and Deep Learning (INNS BDDL) international conference, a major event for researchers in the field of artificial neural networks, big data and related topics, organized by the International Neural Network Society and hosted by the University of Genoa. In 2019 INNS BDDL has been held in Sestri Levante (Italy) from April 16 to April 18. More than 80 researchers from 20 countries participated in the INNS BDDL in April 2019. In addition to regular sessions, INNS BDDL welcomed around 40 oral communications, 6 tutorials have been presented together with 4 invited plenary speakers. This book covers a broad range of topics in big data and deep learning, from theoretical aspects to state-of-the-art applications. This book is directed to both Ph.D. students and Researchers in the field in order to provide a general picture of the state-of-the-art on

the topics addressed by the conference.

Get to know the 'why' and 'how' of machine learning and big data in quantitative investment Big Data and Machine Learning in Quantitative Investment is not just about demonstrating the maths or the coding. Instead, it's a book by practitioners for practitioners, covering the questions of why and how of applying machine learning and big data to quantitative finance. The book is split into 13 chapters, each of which is written by a different author on a specific case. The chapters are ordered according to the level of complexity; beginning with the big picture and taxonomy, moving onto practical applications of machine learning and finally finishing with innovative approaches using deep learning. • Gain a solid reason to use machine learning • Frame your question using financial markets laws • Know your data • Understand how machine learning is becoming ever more sophisticated Machine learning and big data are not a magical solution, but appropriately applied, they are extremely effective tools for quantitative investment — and this book shows you how.

Publisher's note: This is a 2nd edition due to an article retraction

This book constitutes revised selected papers from the First International Workshop on Machine Learning, Optimization, and Big Data, MOD 2015, held in Taormina, Sicily, Italy, in July 2015. The 32 papers presented in this volume were carefully reviewed and selected from 73 submissions. They deal with the algorithms, methods and theories relevant in data science, optimization and machine learning.

This edited volume is devoted to Big Data Analysis from a Machine Learning standpoint as presented by some of the most eminent researchers in this area. It demonstrates that Big Data Analysis opens up new research problems which were either never considered before, or were only considered within a limited range. In addition to providing methodological discussions on the principles of mining Big Data and the difference between traditional statistical data analysis and newer computing frameworks, this book presents recently developed algorithms affecting such areas as business, financial forecasting, human mobility, the Internet of Things, information networks, bioinformatics, medical systems and life science. It explores, through a number of specific examples, how the study of Big Data Analysis has evolved and how it has started and will most likely continue to affect society. While the benefits brought upon by Big Data Analysis are underlined, the book also discusses some of the warnings that have been issued concerning the potential dangers of Big Data Analysis along with its pitfalls and challenges.

Cet ouvrage s'adresse à tous ceux qui cherchent à tirer parti de l'énorme potentiel des « technologies Big Data », qu'ils soient data scientists, DSI, chefs de projets ou spécialistes métier. Le Big Data s'est imposé comme une innovation majeure pour toutes les entreprises qui cherchent à construire un avantage concurrentiel grâce à l'exploitation de leurs données clients, fournisseurs, produits, processus, machines, etc. Mais quelle solution technique choisir ? Quelles compétences métier développer au sein de la DSI ? Ce livre est un guide pour comprendre les enjeux d'un projet Big Data, en appréhender les concepts sous-jacents (en particulier le Machine Learning) et acquérir les compétences nécessaires à la mise en place d'un data lab. Il combine la présentation : • de notions théoriques (traitement statistique des données, calcul distribué...) ; • des outils les plus répandus (écosystème Hadoop, Storm...) ; • d'exemples d'applications ; • d'une organisation typique d'un projet de data science. Les ajouts de cette troisième édition concernent principalement la vision d'architecture d'entreprise, nécessaire pour intégrer les innovations du Big Data au sein des organisations, et le Deep Learning pour le NLP (Natural Language Processing, qui est l'un des domaines de l'intelligence artificielle qui a le plus progressé récemment).

The Era of Artificial Intelligence, Machine Learning and Data Science in the Pharmaceutical Industry examines the drug discovery process, assessing how new technologies have improved effectiveness. Artificial intelligence and machine learning are considered the future for a wide range of disciplines and industries, including the pharmaceutical industry. In an environment where producing a single approved drug costs millions and takes many years of rigorous testing prior to its approval, reducing costs and time is of high interest. This book follows the journey that a drug company takes when producing a therapeutic, from the very beginning to ultimately benefitting a patient's life. This comprehensive resource will be useful to those working in the pharmaceutical industry, but will also be of interest to anyone doing research in chemical biology, computational chemistry, medicinal chemistry and bioinformatics. Demonstrates how the prediction of toxic effects is performed, how to reduce costs in testing compounds, and its use in animal research Written by the industrial teams who are conducting the work, showcasing how the technology has improved and where it should be further improved Targets materials for a better understanding of techniques from different disciplines, thus creating a complete guide

A hands-on approach to tasks and techniques in data stream mining and real-time analytics, with examples in MOA, a popular freely available open-source software framework. Today many information sources—including sensor networks, financial markets, social networks, and healthcare monitoring—are so-called data streams, arriving sequentially and at high speed. Analysis must take place in real time, with partial data and without the capacity to store the entire data set. This book presents algorithms and techniques used in data stream mining and real-time analytics. Taking a hands-on approach, the book demonstrates the techniques using MOA (Massive Online Analysis), a popular, freely available open-source software framework, allowing readers to try out the techniques after reading the explanations. The book first offers a brief introduction to the topic, covering big data mining, basic methodologies for mining data streams, and a simple example of MOA. More detailed discussions follow, with chapters on sketching techniques, change, classification, ensemble methods, regression, clustering, and frequent pattern mining. Most of these chapters include exercises, an MOA-based lab session, or both. Finally, the book discusses the MOA software, covering the MOA graphical user interface, the command line, use of its API, and the development of new methods within MOA. The book will be an essential reference for readers who want to use data stream mining as a tool, researchers in innovation or data stream mining, and programmers who want to create new algorithms for MOA.

Master the approaches and principles of Artificial Intelligence (AI) algorithms, and apply them to Data Science projects with Python and Julia code. Aspiring and practicing Data

Science and AI professionals, along with Python and Julia programmers, will practice numerous AI algorithms and develop a more holistic understanding of the field of AI, and will learn when to use each framework to tackle projects in our increasingly complex world. The first two chapters introduce the field, with Chapter 1 surveying Deep Learning models and Chapter 2 providing an overview of algorithms beyond Deep Learning, including Optimization, Fuzzy Logic, and Artificial Creativity. The next chapters focus on AI frameworks; they contain data and Python and Julia code in a provided Docker, so you can practice. Chapter 3 covers Apache's MXNet, Chapter 4 covers TensorFlow, and Chapter 5 investigates Keras. After covering these Deep Learning frameworks, we explore a series of optimization frameworks, with Chapter 6 covering Particle Swarm Optimization (PSO), Chapter 7 on Genetic Algorithms (GAs), and Chapter 8 discussing Simulated Annealing (SA). Chapter 9 begins our exploration of advanced AI methods, by covering Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). Chapter 10 discusses optimization ensembles and how they can add value to the Data Science pipeline. Chapter 11 contains several alternative AI frameworks including Extreme Learning Machines (ELMs), Capsule Networks (CapsNets), and Fuzzy Inference Systems (FIS). Chapter 12 covers other considerations complementary to the AI topics covered, including Big Data concepts, Data Science specialization areas, and useful data resources to experiment on. A comprehensive glossary is included, as well as a series of appendices covering Transfer Learning, Reinforcement Learning, Autoencoder Systems, and Generative Adversarial Networks. There is also an appendix on the business aspects of AI in data science projects, and an appendix on how to use the Docker image to access the book's data and code. The field of AI is vast, and can be overwhelming for the newcomer to approach. This book will arm you with a solid understanding of the field, plus inspire you to explore further.

This volume comprises six well-versed contributed chapters devoted to report the latest fi ndings on the applications of machine learning for big data analytics. Big data is a term for data sets that are so large or complex that traditional data processing application software is inadequate to deal with them. The possible challenges in this direction include capture, storage, analysis, data curation, search, sharing, transfer, visualization, querying, updating and information privacy. Big data analytics is the process of examining large and varied data sets - i.e., big data - to uncover hidden patterns, unknown correlations, market trends, customer preferences and other useful information that can help organizations make more-informed business decisions. This volume is intended to be used as a reference by undergraduate and post graduate students of the disciplines of computer science, electronics and telecommunication, information science and electrical engineering. THE SERIES: FRONTIERS IN COMPUTATIONAL INTELLIGENCE The series Frontiers In Computational Intelligence is envisioned to provide comprehensive coverage and understanding of cutting edge research in computational intelligence. It intends to augment the scholarly discourse on all topics relating to the advances in artifi cial life and machine learning in the form of metaheuristics, approximate reasoning, and robotics. Latest research fi ndings are coupled with applications to varied domains of engineering and computer sciences. This field is steadily growing especially with the advent of novel machine learning algorithms being applied to different domains of engineering and technology. The series brings together leading researchers that intend to continue to advance the fi eld and create a broad knowledge about the most recent research.

Cet ouvrage s'adresse à tous ceux qui réfléchissent à la meilleure utilisation possible des données au sein de l'entreprise, qu'ils soient data scientists, DSI, chefs de projets ou spécialistes métier. Le Big Data s'est imposé comme une innovation majeure pour toutes les entreprises qui cherchent à construire un avantage concurrentiel grâce à l'exploitation de leurs données clients, fournisseurs, produits, processus, machines, etc. Mais quelle solution technique choisir ? Quelles compétences métier développer au sein de la DSI ? Ce livre est un guide pour comprendre les enjeux d'un projet Big Data, en appréhender les concepts sous-jacents (en particulier le machine learning) et acquérir les compétences nécessaires à la mise en place d'un data lab. Il combine la présentation : • de notions théoriques (traitement statistique des données, calcul distribué...) ; • d'outils (écosystème Hadoop, Storm...) ; • d'exemples de machine learning ; • d'une organisation typique d'un projet de data science.

This book presents deep learning techniques, concepts, and algorithms to classify and analyze big data. Further, it offers an introductory level understanding of the new programming languages and tools used to analyze big data in real-time, such as Hadoop, SPARK, and GRAPHX. Big data analytics using traditional techniques face various challenges, such as fast, accurate and efficient processing of big data in real-time. In addition, the Internet of Things is progressively increasing in various fields, like smart cities, smart homes, and e-health. As the enormous number of connected devices generate huge amounts of data every day, we need sophisticated algorithms to deal, organize, and classify this data in less processing time and space. Similarly, existing techniques and algorithms for deep learning in big data field have several advantages thanks to the two main branches of the deep learning, i.e. convolution and deep belief networks. This book offers insights into these techniques and applications based on these two types of deep learning. Further, it helps students, researchers, and newcomers understand big data analytics based on deep learning approaches. It also discusses various machine learning techniques in concatenation with the deep learning paradigm to support high-end data processing, data classifications, and real-time data processing issues. The classification and presentation are kept quite simple to help the readers and students grasp the basics concepts of various deep learning paradigms and frameworks. It mainly focuses on theory rather than the mathematical background of the deep learning concepts. The book consists of 5 chapters, beginning with an introductory explanation of big data and deep learning techniques, followed by integration of big data and deep learning techniques and lastly the future directions.

Build next-generation Artificial Intelligence systems with Java Key Features Implement AI techniques to build smart applications using Deeplearning4j Perform big data analytics to derive quality insights using Spark MLlib Create self-learning systems using neural networks, NLP, and reinforcement learning Book Description In this age of big data,

companies have larger amount of consumer data than ever before, far more than what the current technologies can ever hope to keep up with. However, Artificial Intelligence closes the gap by moving past human limitations in order to analyze data. With the help of Artificial Intelligence for big data, you will learn to use Machine Learning algorithms such as k-means, SVM, RBF, and regression to perform advanced data analysis. You will understand the current status of Machine and Deep Learning techniques to work on Genetic and Neuro-Fuzzy algorithms. In addition, you will explore how to develop Artificial Intelligence algorithms to learn from data, why they are necessary, and how they can help solve real-world problems. By the end of this book, you'll have learned how to implement various Artificial Intelligence algorithms for your big data systems and integrate them into your product offerings such as reinforcement learning, natural language processing, image recognition, genetic algorithms, and fuzzy logic systems. What you will learn Manage Artificial Intelligence techniques for big data with Java Build smart systems to analyze data for enhanced customer experience Learn to use Artificial Intelligence frameworks for big data Understand complex problems with algorithms and Neuro-Fuzzy systems Design stratagems to leverage data using Machine Learning process Apply Deep Learning techniques to prepare data for modeling Construct models that learn from data using open source tools Analyze big data problems using scalable Machine Learning algorithms Who this book is for This book is for you if you are a data scientist, big data professional, or novice who has basic knowledge of big data and wish to get proficiency in Artificial Intelligence techniques for big data. Some competence in mathematics is an added advantage in the field of elementary linear algebra and calculus. Present book covers new paradigms in Blockchain, Big Data and Machine Learning concepts including applications and case studies. It explains dead fusion in realizing the privacy and security of blockchain based data analytic environment. Recent research of security based on big data, blockchain and machine learning has been explained through actual work by practitioners and researchers, including their technical evaluation and comparison with existing technologies. The theoretical background and experimental case studies related to real-time environment are covered as well. Aimed at Senior undergraduate students, researchers and professionals in computer science and engineering and electrical engineering, this book: Converges Blockchain, Big Data and Machine learning in one volume. Connects Blockchain technologies with the data centric applications such Big data and E-Health. Easy to understand examples on how to create your own blockchain supported by case studies of blockchain in different industries. Covers big data analytics examples using R. Includes Illustrative examples in python for blockchain creation. Publisher's Note: Products purchased from Third Party sellers are not guaranteed by the publisher for quality, authenticity, or access to any online entitlements included with the product. Use machine learning to understand your customers, frame decisions, and drive value The business analytics world has changed, and Data Scientists are taking over. Business Data Science takes you through the steps of using machine learning to implement best-in-class business data science. Whether you are a business leader with a desire to go deep on data, or an engineer who wants to learn how to apply Machine Learning to business problems, you'll find the information, insight, and tools you need to flourish in today's data-driven economy. You'll learn how to: •Use the key building blocks of Machine Learning: sparse regularization, out-of-sample validation, and latent factor and topic modeling•Understand how use ML tools in •real world business problems, where causation matters more that correlation•Solve data science programs by scripting in the R programming language Today's business landscape is driven by data and constantly shifting. Companies live and die on their ability to make and implement the right decisions quickly and effectively. Business Data Science is about doing data science right. It's about the exciting things being done around Big Data to run a flourishing business. It's about the precepts, principals, and best practices that you need know for best-in-class business data science. Summary Introducing Data Science teaches you how to accomplish the fundamental tasks that occupy data scientists. Using the Python language and common Python libraries, you'll experience firsthand the challenges of dealing with data at scale and gain a solid foundation in data science. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Many companies need developers with data science skills to work on projects ranging from social media marketing to machine learning. Discovering what you need to learn to begin a career as a data scientist can seem bewildering. This book is designed to help you get started. About the Book Introducing Data ScienceIntroducing Data Science explains vital data science concepts and teaches you how to accomplish the fundamental tasks that occupy data scientists. You'll explore data visualization, graph databases, the use of NoSQL, and the data science process. You'll use the Python language and common Python libraries as you experience firsthand the challenges of dealing with data at scale. Discover how Python allows you to gain insights from data sets so big that they need to be stored on multiple machines, or from data moving so quickly that no single machine can handle it. This book gives you hands-on experience with the most popular Python data science libraries, Scikit-learn and StatsModels. After reading this book, you'll have the solid foundation you need to start a career in data science. What's Inside Handling large data Introduction to machine learning Using Python to work with data Writing data science algorithms About the Reader This book assumes you're comfortable reading code in Python or a similar language, such as C, Ruby, or JavaScript. No prior experience with data science is required. About the Authors Davy Cielen, Arno D. B. Meysman, and Mohamed Ali are the founders and managing partners of Optimately and Maiton, where they focus on developing data science projects and solutions in various sectors. Table of Contents Data science in a big data world The data science process Machine learning Handling large data on a single computer First steps in big data Join the NoSQL movement The rise of graph databases Text mining and text analytics Data visualization to the end user Interest in big data has swelled within the scholarly community as has increased attention to the internet of things (IoT). Algorithms are constructed in order to parse and analyze all this data to facilitate the exchange of information. However, big data has suffered from problems in connectivity, scalability, and privacy since its birth. The application of deep learning algorithms has helped process those challenges and remains a major issue in today's digital world. Advanced Deep Learning Applications in Big Data Analytics is a pivotal reference source that aims to develop new architecture and applications of deep learning algorithms in big data and the IoT. Highlighting a wide range of topics such as artificial intelligence, cloud computing, and neural networks, this book is ideally designed for engineers, data analysts, data scientists, IT specialists, programmers, marketers, entrepreneurs, researchers, academicians, and students.